

BAB 3

PELAKSANAAN KERJA MAGANG

3.1 Kedudukan dan Organisasi

Pelaksanaan kerja magang di Bappeda Pohuwato, disalurkan ke perusahaan mitra bernama Tatiye. Posisi yang diberikan adalah sebagai ilmuwan data. Pelaksana kerja magang diposisikan di tim pemegang yang berfokus pada pengujian dan pemanfaatan Portal Satu Data Indoneisa-Pohuwato sebagai sebuah portal data. Tugas yang diberikan kepada pelaksana kerja magang sebagai ilmuwan data meliputi pembuatan API dokumentasi dan *fine-tuning* model *chatbot* dengan menyesuaikan permintaan dan kebutuhan pengguna. Selain itu, ilmuwan data juga bertanggung jawab terhadap pemeliharaan kualitas data yang ada di portal data. Supervisi memantau perkembangan semua proyek yang diberikan secara berkala di tempat kerja. *Briefing* dilakukan setiap pagi untuk mengevaluasi pekerjaan dihari sebelumnya, kemudian diberikan target baru untuk hari itu.

3.2 Tugas yang Dilakukan

Selama pelaksanaan kerja magang di Badan Perencanaan Pembangunan Daerah Kabupaten Pohuwato, tugas yang diberikan adalah pembuatan pembuatan API dokumentasi dan *fine-tuning* model *chatbot*. *Fine-tuning* bertujuan untuk mengintegrasikan data yang ada di Posdip agar mampu membantu pengguna dalam lingkup pemerintah Kabupaten Pohuwato lebih mudah dalam mengakses data. Model *chatbot* dikembangkan dengan data yang beragam sesuai yang ada di aplikasi Open Data Pohuwato.

Model *chatbot* dikembangkan dengan menggunakan bahasa pemrograman Python. Alasan menggunakan bahasa pemrograman Python adalah bahasa pemrograman yang open source, fleksibel, dan ramah bagi pemula. Selain itu, banyak model lain sudah dikembangkan lewat bahasa pemograman Python, sehingga pengaplikasiannya lebih mudah. Library utama yang digunakan adalah `google.generativeai` untuk memuat model dari Gemini. Alasan menggunakan model ini adalah karena dapat diakses secara gratis, memiliki pemrosesan Bahasa Indonesia yang baik, dan cukup ringan dibanding melatih model lainnya.

Dalam pelaksanaan *fine-tuning*, capaian yang diinginkan adalah model mampu memberikan jawaban yang sesuai dengan lingkup pengetahuan tentang aplikasi

Posdip. Model mampu memberikan penjelasan tentang fitur yang ada pada aplikasi Posdip. Kemudian model mampu memberikan tampilan dan analisis sederhana yang sesuai dengan data yang ada pada aplikasi Posdip.

3.3 Uraian Pelaksanaan Magang

Pelaksanaan kerja magang di Badan Perencanaan Pembangunan Daerah Kabupaten Pohuwat untuk pelatihan model gemini untuk kebutuhan Pos Satu Data Indonesia-Pohuwato diuraikan seperti pada Tabel 3.1.

Tabel 3.1. Pekerjaan yang dilakukan tiap minggu selama pelaksanaan kerja magang

Minggu Ke -	Pekerjaan yang dilakukan.
1	Pengenalan lingkungan kerja, <i>jobdesk</i> , dan proyek yang akan dikerjakan.
2	Melakukan proses pra-produksi yang terdiri dari riset untuk pemilihan dan penggunaan model, mengeksplorasi data, dan memahami cara kerja aplikasi Posdip.
3	Mempelajari library yang digunakan untuk melakukan pemodelan dan pelatihan, seperti <i>google.generative</i> , <i>os</i> , <i>json</i> , dan <i>pandas</i> .
4	Melakukan ekstraksi dan <i>preprocessing data</i> .
5	Memulai <i>fine-tuning</i> tahap awal terkait informasi dasar aplikasi dan daerah.
6	Melanjutkan <i>fine-tuning</i> dengan menggunakan data pemerintahan yang tersedia.
7	melakukan <i>Update</i> data yang tersedia pada aplikasi Posdip.
8	Mengimplementasikan model secara lokal.
9	Melakukan optimasi <i>prompt</i> untuk meringankan penggunaan sumber daya.
10	Melakukan tes terhadap model.
11	Melakukan riset untuk menerapkan metode <i>retrieval augmented generation</i> .
12	Finalisasi model.

3.3.1 User Requirement dan perangkat penunjang

Dalam pengembangan asisten untuk keperluan penyajian dan analisis data dalam lingkup pemerintahan, terdapat beberapa requirement yang perlu diperhatikan.

1. Model mampu memproses *prompt* dengan bahasa yang sesuai dengan kebiasaan masyarakat.
2. Model mampu memberikan data dan analisis sederhana dengan data yang ada di aplikasi Posdip.

Dalam pengembangannya, terdapat beberapa perangkat penunjang untuk pemrosesan data dan pengembangan model.

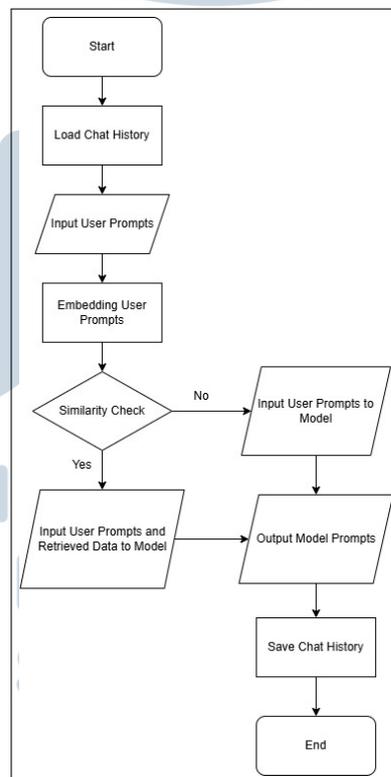
1. Python: bahasa pemrograman Python adalah pilihan utama dalam pemrosesan data dan pengembangan model, karena kesederhanaannya dan ekosistemnya yang kaya. Python mendukung berbagai *library*, seperti *pandas* dan *json* untuk bagian data, *google.generativeai* sebagai model *chatbot*, *sentence transformers* berguna untuk *embedding*, dan *faiss* sebagai *search vector*.
2. *pandas*: adalah *library* Python yang dirancang untuk pemrosesan data. Menyediakan struktur data fleksibel seperti *DataFrame* (struktur dua dimensi seperti tabel) untuk mengelola data secara efisien. *Pandas* memungkinkan pembacaan dan penulisan data dalam berbagai format, seperti *CSV*, *Excel*, *JSON*, dan *SQL*.
3. *json*: adalah *library* Python yang digunakan dengan tujuan untuk bekerja dengan data *JSON* (*JavaScript Object Notation*), mengonversi data Python seperti *dictionary* atau *list* menjadi string *JSON*, dan mengonversi data *JSON* menjadi struktur data Python.
4. *googleGenerativeai*: adalah *library* Python yang dirancang untuk berinteraksi dengan model bahasa besar (*Large Language Models, LLM*) dari Google, seperti model *Gemini*, melalui *API Generative AI*. Pustaka ini dirancang agar mudah diintegrasikan dengan proyek Python yang membutuhkan solusi *NLP* berbasis *AI* generatif.
5. *sentenceTransformers*: adalah pustaka Python berbasis model *transformer* yang dirancang untuk menghasilkan representasi vektor (*embeddings*)

dari teks, seperti kalimat atau paragraf, sehingga memungkinkan analisis kesamaan semantik, pencarian informasi, clustering, dan tugas Natural Language Processing (NLP) lainnya.

6. *faiss (Facebook AI Similarity Search)*: adalah *library* yang dikembangkan oleh Meta AI untuk memproses pencarian kesamaan berbasis vektor dengan efisiensi tinggi, terutama dalam skenario dengan dataset besar. *Faiss* dirancang untuk menangani tugas-tugas seperti pencocokan, pengambilan informasi, dan pengelompokan data vektor dalam dimensi tinggi.

3.3.2 Perancangan Sistem

Perancangan model *assistant* terdiri dari empat tahapan, yaitu perancangan sistem, koleksi dan pemrosesan data, pelatihan model, dan penerapan model. Tahapan yang dilakukan pada tahap perancangan sistem berfokus untuk melakukan riset serta pembuatan *flowchart*. Tujuan dari pembuatan *flowchart* adalah untuk mengetahui alur dari cara kerja model chatbot yang akan dibuat. *Flowchart* dari *assistant* yang akan diimplementasikan di Aplikasi *Posdip* dapat dilihat pada gambar 3.1



Gambar 3.1. *Flowchart* model *assistant* pada *Open Data*

3.3.3 Pemrosesan dan koleksi data

Bagian ini memuat metode untuk mengumpulkan dan memroses data agar sesuai dengan kebutuhan model yang akan di *fine tune*. Bentuk data yang dikumpulkan adalah data pemerintahan dari berbagai OPD di Kabupaten Pohuwato. Format data yang dikumpulkan dalam format excel dan JSON, hal ini dilakukan dengan melihat kemampuan aplikasi Posdip dalam mengkonversi data yang diinginkan. Aplikasi Posdip sebenarnya mampu untuk memberikan data dalam format seperti CSV (*Comma-Separated Values*) dan PDF (*Portable Document Format*), namun karena terkendala pada kualitas data yang ada, dipilihlah format JSON dan excel.

The screenshot shows a web interface for an Open Data portal. The title is "Persentase Penduduk Menurut Kecamatan dan Agama yang Dianut di Kabupaten Pohuwato (Persen)". The source is "Badan Pusat Statistik Indonesia". The table below displays the data for 2019 and 2020, categorized by district and religion.

Kecamatan dan Kabupaten	2019						2020					
	Islam	Protestan	Katolik	Hindu	Buddha	Konghucu	Islam	Protestan	Katolik	Hindu	Buddha	Konghucu
Popayato	96.68	96.91	3.21	2.95	0.04	0.04	0.07	0.11	0	0	0	0
Popayato Barat	98.29	98.29	1.44	1.44	0.22	0.22	0.05	0.05	0	0	0	0
Popayato Timur	75.36	76.8	24.1	23.2	0.48	0	0.06	0	0	0	0	0
Lemito	97.37	97.37	2.63	2.63	0	0	0	0	0	0	0	0
Wanggarasi	94.56	94.56	5.3	5.3	0.14	0.14	0	0	0	0	0	0
Marisa	99.08	99.08	0.78	0.78	0.04	0.04	0.1	0.1	0	0	0	0
Patilanggio	98.37	98.2	1.63	1.8	0	0	0	0	0	0	0	0
Buntulia	99.66	99.66	0.29	0.29	0.03	0.03	0	0	0.02	0.02	0	0
Duhiadaa	98.91	98.91	1.04	1.04	0.01	0.01	0	0	0.03	0.03	0	0
Randangan	90	90	2.97	2.97	0.49	0.49	6.48	6.48	0.06	0.06	0	0
Taluditi	92.85	92.85	5.66	5.66	0.25	0.25	1.23	1.23	0	0	0	0
Paguat	98.66	98.66	1.25	1.25	0.06	0.06	0.01	0.01	0.02	0.02	0	0
Dengilo	90.94	90.94	9.06	9.06	0	0	0	0	0	0	0	0

Gambar 3.2. Tampilan tabel data yang ada di *Open Data*

Setelah berhasil mendapatkan seluruh data yang ada pada aplikasi Posdip, data diklasifikasi menjadi beberapa kategori berdasarkan OPD dimana data tersebut dirilis. Terdapat 130 data yang dibagi menjadi 10 kategori, yaitu sosial, pendidikan, kesehatan, kependudukan, ekonomi, pemerintahan, pariwisata, lingkungan hidup, kemiskinan, dan infrastruktur. Persebaran jumlah data tidak merata pada setiap kategori, sosial memiliki 5 data, pendidikan memiliki 6 data, kesehatan memiliki 15 data, kependudukan memiliki 24 data, ekonomi memiliki 66 data, pemerintahan memiliki 3 data, pariwisata memiliki 1 data, lingkungan hidup memiliki 6 data, kemiskinan memiliki 1 data, dan infrastruktur memiliki 3 data.

	2019										2020		
	Islam	Protestan	Katolik	Hindu	Buddha	Konghucu	Islam	Protestan	Katolik	Hindu	Buddha	Konghucu	
Popayato	96,68	96,91	3,21	2,95	0,04	0,04	0,07	0,11	0	0	0	0	0
Popayato	98,29	98,29	1,44	1,44	0,22	0,22	0,05	0,05	0	0	0	0	0
Popayato	75,36	76,8	24,1	23,2	0,48	0	0,06	0	0	0	0	0	0
Kecamatan Lemito	97,37	97,37	2,63	2,63	0	0	0	0	0	0	0	0	0
Wanggara	94,56	94,56	5,3	5,3	0,14	0,14	0	0	0	0	0	0	0
Marisa	99,08	99,08	0,78	0,78	0,04	0,04	0,1	0,1	0	0	0	0	0
Patilanggi	98,37	98,2	1,63	1,8	0	0	0	0	0	0	0	0	0
Buntulia	99,66	99,66	0,29	0,29	0,03	0,03	0	0	0,02	0,02	0	0	0
Duhiadaa	98,91	98,91	1,04	1,04	0,01	0,01	0	0	0,03	0,03	0	0	0
Randanga	90	90	2,97	2,97	0,49	0,49	6,48	6,48	0,06	0,06	0	0	0
Taluditi	92,85	92,85	5,66	5,66	0,25	0,25	1,23	1,23	0	0	0	0	0
Paguat	98,66	98,66	1,25	1,25	0,06	0,06	0,01	0,01	0,02	0,02	0	0	0
Dengilo	90,94	90,94	9,06	9,06	0	0	0	0	0	0	0	0	0
Pohuwato	95,27	95,33	3,69	3,66	0,14	0,11	0,89	0,89	0,01	0,01	0	0	0

Gambar 3.3. Tampilan data excel yang diambil dari *Open Data*



```
[
{
  "2019": "Protestan",
  "2020": "Katolik",
  "Kecamatan dan Kabupaten": "Islam"
},
{
  "2019": "96.68",
  "2020": "96.91",
  "Kecamatan dan Kabupaten": "Popayato"
},
{
  "2019": "98.29",
  "2020": "98.29",
  "Kecamatan dan Kabupaten": "Popayato Barat"
},
{
  "2019": "75.36",
  "2020": "76.8",
  "Kecamatan dan Kabupaten": "Popayato Timur"
},
{
  "2019": "97.37",
  "2020": "97.37",
  "Kecamatan dan Kabupaten": "Lemito"
},
]
```

Gambar 3.4. Tampilan data JSON yang diambil dari *Open Data*

Melihat kesalahan yang ada pada gambar 3.3 dan 3.4, kesalahan format data terjadi pada data yang memiliki sub kolom. Dengan begitu, perbaikan terhadap baris dan kolom data excel yang rusak dikerjakan secara manual. Kemudian Data dalam format excel yang telah diperbaiki dapat dilihat pada gambar 3.5. Sedangkan data JSON diperbaiki dan disatukan menjadi *dictionary* dengan menggunakan program yang dibuat dalam bahasa pemrograman Python. Format data JSON yang sudah

diperbaiki dapat dilihat pada gambar 3.6. Setelah memiliki data yang mumpuni, dilanjutkan dengan pembuatan *prompt input* dan *output* untuk menunjang proses pelatihan model.

Kecamatan dan Kabupaten	TK	PAUD	SD	MI	SMP	MTs	SMA	SMK	MA
Popayato	4	1	0	1	2	1	1	1	0
Popayato Barat	4	2	0	1	2	0	0	1	0
Popayato Timur	3	0	0	0	0	0	0	0	0
Lemito	4	1	0	1	2	0	0	1	0
Wanggarasi	3	0	0	0	1	1	1	0	0
Marisa	6	2	0	1	3	2	1	1	0
Patilanggio	3	1	0	0	1	1	1	1	0
Buntulia	4	1	0	1	2	0	0	1	0
Duhiadaa	3	0	0	0	0	1	1	0	0
Randangan	6	2	0	1	2	2	2	1	0
Taluditi	5	0	0	0	0	1	1	0	0
Paguat	5	2	0	1	2	0	0	1	0
Dengilo	2	0	0	0	0	1	1	0	0
Pohuwato	52	0	0	0	17	0	9	0	0

Gambar 3.5. Tampilan data excel dari *Open Data* dengan format yang telah disesuaikan



```
{
  "category": "ekonomi",
  "title": "Distribusi PDRB Atas Dasar Harga Berlaku menurut Lapangan Usaha (Persen).xlsx",
  "data": [
    {
      "Lapangan Usaha": "Pertanian, Kehutanan, dan Perikanan",
      "2010": 58.38,
      "2011": 58.93,
      "2012": 58.97,
      "2013": 59.03,
      "2014": 59.57,
      "2015": 59.07,
      "2016": 59.22,
      "2017": 59.75,
      "2018": 60.19,
      "2019": 60.29,
      "2020": 59.9,
      "2021": 59.97,
      "2022": 58.49,
      "2023": 57.35
    },
    {
      "Lapangan Usaha": "Pertambangan dan Penggalian",
      "2010": 1.45,
      "2011": 1.41,
      "2012": 1.29,
      "2013": 1.15,
      "2014": 1.03,
      "2015": 0.94,
      "2016": 0.82,
      "2017": 0.76,
      "2018": 0.71,
      "2019": 0.71,
      "2020": 0.77,
      "2021": 0.77,
      "2022": 0.78,
      "2023": 0.76
    }
  ]
}
```

Gambar 3.6. Tampilan data JSON dari *Open Data* dengan format yang telah disesuaikan

3.3.4 Mempersiapkan Data Prompt

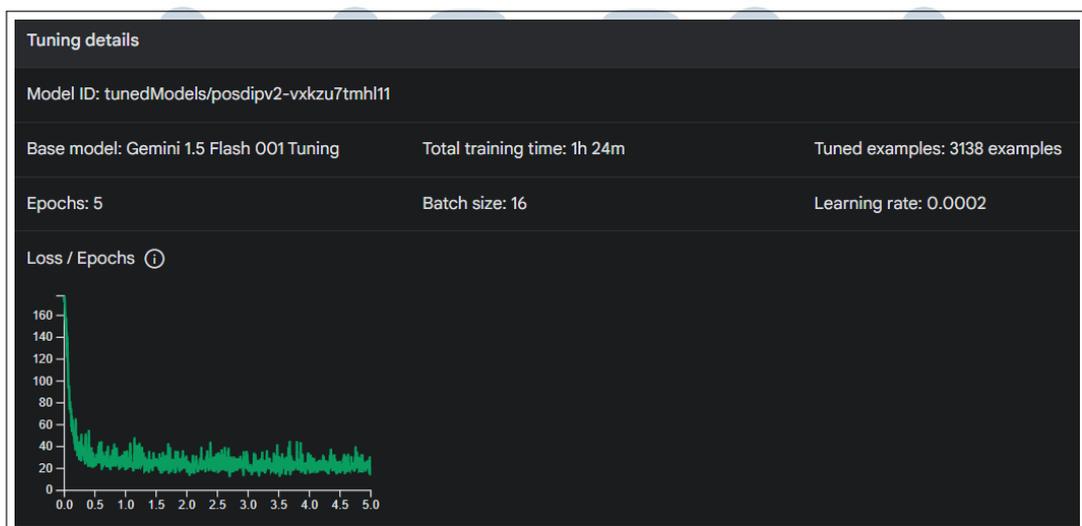
Untuk proses *fine tuning* menggunakan model dasar Gemini, dibutuhkan *prompt* dan ekspektasi output yang akan diberikan oleh model. Maka untuk mempersiapkan data *prompt* yang dibutuhkan, proses pembuatan *prompt* dibagi menjadi dalam 2 tahap, yaitu menulis langsung *prompt* untuk menambahkan informasi umum seputar aplikasi Posdip dan fitur yang ada didalamnya. Tahap lainnya adalah merubah data yang sudah diproses sebelumnya menjadi *prompt* dengan menulis seluruh *key* dari setiap *attribute* dan merubah nilainya sebagai *expected outputnya*. Semua *prompt* yang telah dibuat akan diubah dalam tabel excel dengan kolom input dan output.

user	model
Apa itu aplikasi POSDIP?	Posdip adalah aplikasi Portal Satu Data Indonesia Puhuwato untuk mengakses seluruh data data skpd yang berac
Bagaimana saya mengakses Posdip?	Untuk mengakses Posdip, kamu bisa mengikuti link ini : https://opendata.puhuwatokab.go.id/ . Sebelum itu pasti
Apakah kamu tahu Tatiye.id?	Tatiye adalah perusahaan IT Solution yang berada di kabupaten puhuwato, merupakan perusahaan yang menge
Apa fungsi dari fitur assistant di aplikasi posdip?	Iya, fungsi assistant ini membantu user untuk memahami aplikasi Posdip dan data-data yang ditampungnya.
Apa fungsi dari file explorer di aplikasi posdip?	Fungsi dari file explorer pada aplikasi Posdip, memungkinkan Anda untuk menjelajahi dan mengelola file yang di
Apa fungsi dari fitur Analytics di aplikasi posdip	Fungsi fitur analytic pada Posdip, yaitu menyediakan analisis data atau laporan tentang penggunaan aplikasi atau
Bagaimana cara mengubah atau memperbarui inform	Untuk memperbarui informasi akun, tekan tombol Setelan di halaman ini. Setelah itu, Anda akan diarahkan ke h
Apa fungsi dari opsi Buat Folder?	Opsi buat folder digunakan untuk membuat folder baru di dalam direktori saat ini. Folder ini dapat berfungsi unt
Apa yang terjadi jika saya memilih Refresh?	jika kamu memilih Refresh pada Posdip, maka akan memperbarui tampilan halaman atau direktori untuk memas
Apa yang dimaksud dengan File Xlsx pada aplikasi Pos	Xlsx pada aplikasi Posdip, memungkinkan Anda untuk membuat file berformat Excel (XLSX), yang biasanya diguna
Apa kegunaan dari File Charts pada aplikasi Posdip?	File Chart pada aplikasi Posdip memungkinkan Anda untuk membuat file berisi grafik atau chart, biasanya untuk i
Bagaimana cara menggunakan File Pptx pada aplikasi	File Pptx pada aplikasi Posdip merujuk pada PowerPoint Text, fitur ini digunakan untuk membuat file presentasi c
Apa kegunaan dari File docs di aplikasi Posdip?	Pada aplikasi Posdip, file docs memiliki kegunaan untuk membuat dokumen teks dalam format Microsoft Word (l
Apa kegunaan dari File PDF pada aplikasi Posdip?	Pada aplikasi Posdip, fitur dari file PDF memberikan dokumen yang tampilan dan formatnya tetap konsisten, terle
Apa fungsi dari Folder Foto di aplikasi posdip?	Folder Foto pada aplikasi Posdip merupakan folder khusus yang digunakan untuk menyimpan file gambar atau fc
Apa fungsi dari folder di video di aplikasi posdip?	Folder Video pada aplikasi Posdip merupakan folder khusus yang digunakan untuk menyimpan file video dalam s
Apa isi dari folder Kependudukan di bagian Apps pada	Folder kependudukan pada aplikasi Posdip merupakan folder yang berisi data atau informasi terkait kependuduk
Apa isi dari folder Ekonomi di bagian Apps pada apli	Folder Ekonomi pada aplikasi Posdip merupakan folder yang menyimpan data mengenai aspek ekonomi, seperti

Gambar 3.7. Tampilan *prompt* yang telah dibuat

3.3.5 Pelatihan Model

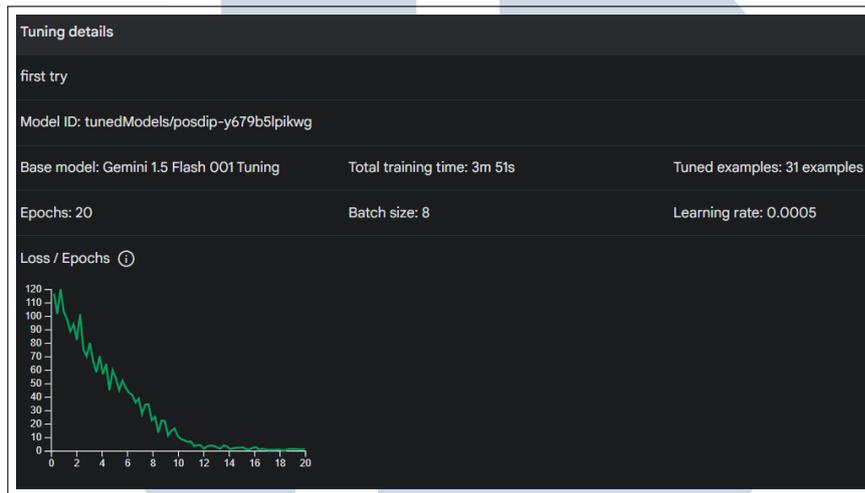
Pelatihan model dilakukan dengan *Google AI Studio*. Setelah memiliki format data yang sesuai, hal yang perlu dipertimbangkan adalah proses dan hasil dari model yang yang berhasil di *fine tune*. Konfigurasi hanya digunakan dibagian menentukan jumlah *epoch*, *batch size*, dan *learning rate*. Dengan begitu model akan dilatih lewat *cloud* sehingga tidak membebani sumber daya internal. Model dasar yang tersedia digunakan untuk *fine tuning* adalah *gemini-1.5-flash-001-tuning*. Dengan begitu hal yang paling berperan adalah kualitas data, konfigurasi *fine tune*, dan evaluasi lewat *plot loss/epoch* yang telah disediakan.



Gambar 3.8. Tampilan dari model yang gagal untuk dilatih

Dikarenakan adanya *underfitting* ketika melatih model dengan data yang ada di aplikasi Posdip. Hanya data *prompt* yang dibuat secara manual yang dimasukkan dalam pelatihan. Hal ini mempertimbangkan kecepatan dan akurasi dari proses

pelatihan model. Namun, dengan begitu maka seluruh data tidak akan masuk dalam pengetahuan dari model. Oleh karena itu pengerjaan dari permasalahan ini akan dijelaskan pada bagian berikutnya dengan metode RAG (*retrieval augmented generation*).



Gambar 3.9. Tampilan dari model yang berhasil untuk dilatih

Dengan adanya masalah pada saat pelatihan model, RAG dipilih untuk membantu model untuk memberikan data yang diminta oleh user lewat *prompt*. Langkah-langkah yang perlu dilakukan adalah membuat *dictionary* yang menyimpan judul data dan nilai dari tabelnya, melakukan *embedding* judul data untuk merubah nilai *string* menjadi vektor, menggunakan *vector searcher* untuk mencari kedekatan dari *prompt* dan judul data yang ada, kemudian mengembalikan untuk nilai tabel dari nama judul yang sesuai. Proyek ini menggunakan sentenceTransformer dengan model all-MiniLM-L6-v2 untuk proses *embedding* text, setelah itu vektor akan disimpan dan digunakan untuk proses pencarian nanti.

```
array([[ -0.04415069,  0.07065102, -0.0756069 , ...,  0.08414089,
        -0.02572981,  0.02388876],
       [-0.06126476,  0.07775955, -0.04941744, ...,  0.08191146,
        -0.05050281,  0.02188062],
       [-0.06921204,  0.11478876,  0.00766862, ...,  0.04921588,
        -0.0358852 ,  0.02285588],
       ...,
       [-0.02643938,  0.11492916, -0.02041236, ...,  0.07872073,
         0.01731305, -0.0212962 ],
       [-0.08369797,  0.0023396 , -0.1094991 , ...,  0.11007672,
         0.09189951, -0.03815973],
       [-0.03882604,  0.07169074, -0.07919402, ..., -0.0139202 ,
         0.01248155,  0.01047403]], shape=(251, 384), dtype=float32)
```

Gambar 3.10. Tampilan vektor dari judul tabel yang telah diproses lewat teknik *embedding*

Setelah data string diubah menjadi vektor, pencarian terhadap vektor bisa dilakukan. Dengan menggunakan *faiss (Facebook AI Similarity Search)*, model bisa melakukan pencarian kesamaan (*similarity search*) dan *clustering* pada vektor berdimensi tinggi. Hal ini berkerja pada setiap *prompt* pengguna yang masuk untuk mencari nilai jarak dari kesamaan *prompt* pengguna dengan vektor hasil *embedding* sebelumnya. *prompt* dari user akan diubah menjadi vektor juga, kemudian *faiss* akan mencari jarak terdekat dengan vektor yang telah disimpan sebelumnya. Hasil dari pencarian vektor dapat di lihat pada gambar 3.11

```
search_result[0]

{'title': 'Jumlah Tempat Peribadatan Menurut Kecamatan di Kabupaten Pohuwato',
 'category': 'sosial',
 'distances': np.float32(0.0),
 'data': [{'Kecamatan dan Kabupaten': 'Popayato',
           'Masjid': 2,
           'Mushola': 2,
           'Gereja Protestant': 2,
           'Gereja Katolik': 1,
           'Pura': 1,
           'Vihara': 1},
          {'Kecamatan dan Kabupaten': 'Popayato Barat',
           'Masjid': 1,
           'Mushola': 1,
           'Gereja Protestant': 1,
           'Gereja Katolik': 0,
           'Pura': 0,
           'Vihara': 0},
          {'Kecamatan dan Kabupaten': 'Popayato Timur',
           'Masjid': 6,
           'Mushola': 6,
           'Gereja Protestant': 8,
           'Gereja Katolik': 2,
           'Pura': 2,
           'Vihara': 2},
```

Gambar 3.11. Tampilan hasil dari *vector search*

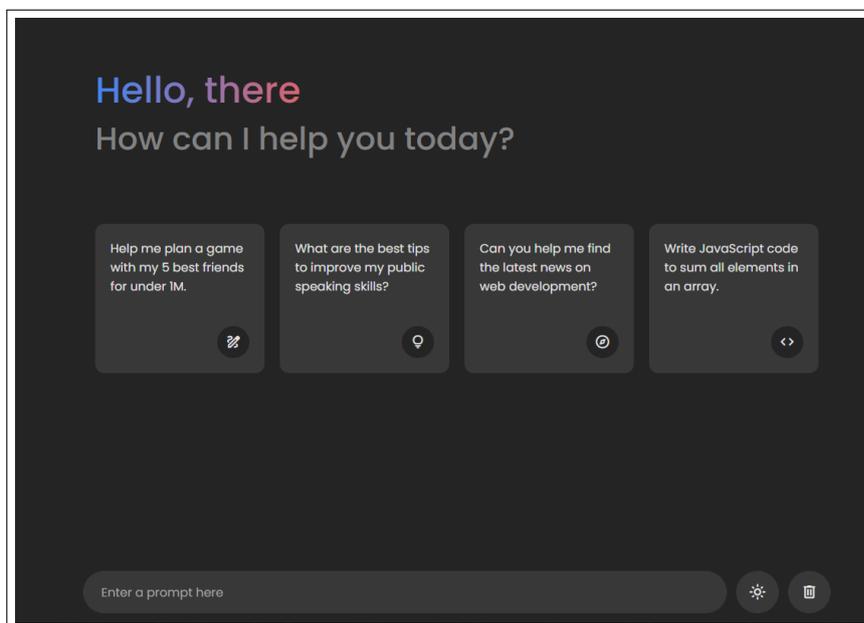
3.3.6 Penerapan Model

Model yang telah di latih pada *Google AI Studio* disimpan dalam *cloud*, sehingga diperlukan pemanggilan nama model yang telah dilatih sesuai dengan kunci *API Application Programming Interface* yang dimiliki. Konfigurasi pada model dapat dikustomisasi dengan parameter yang telah ditentukan pada model tersebut. Kustomisasi dilakukan pada temperatur dan output token maksimal. Hal ini didasarkan untuk menghemat penggunaan sumber daya token yang dimiliki dan menjaga chat yang dihasilkan tetap baku supaya mudah dimengerti oleh pengguna.

```
self.model = genai.GenerativeModel
    model_name=model_name,
    generation_config={
        "temperature": 0.7,
        "top_p": 0.95,
        "max_output_tokens": 8000,
    }
)
```

Gambar 3.12. Tampilan dari konfigurasi model yang digunakan

Penerepan chatbot menggunakan html, css, dan java script sebagai penyokong utama website dan flask sebagai *framework back end*. Pemilihan dari penggunaan alat tersebut dilatar belakangi dengan kemudahan implementasinya nanti ke aplikasi Posdip yang memiliki bahasa pemograman yang serupa. Hasil dari ini juga sudah dilakukan *testing* dan sejauh ini belum ada kendala yang terkait kinerja model pada tahap ini.



Gambar 3.13. Tampilan dari chatbot yang berhasil diimplementasikan

3.3.7 Kendala dan Solusi

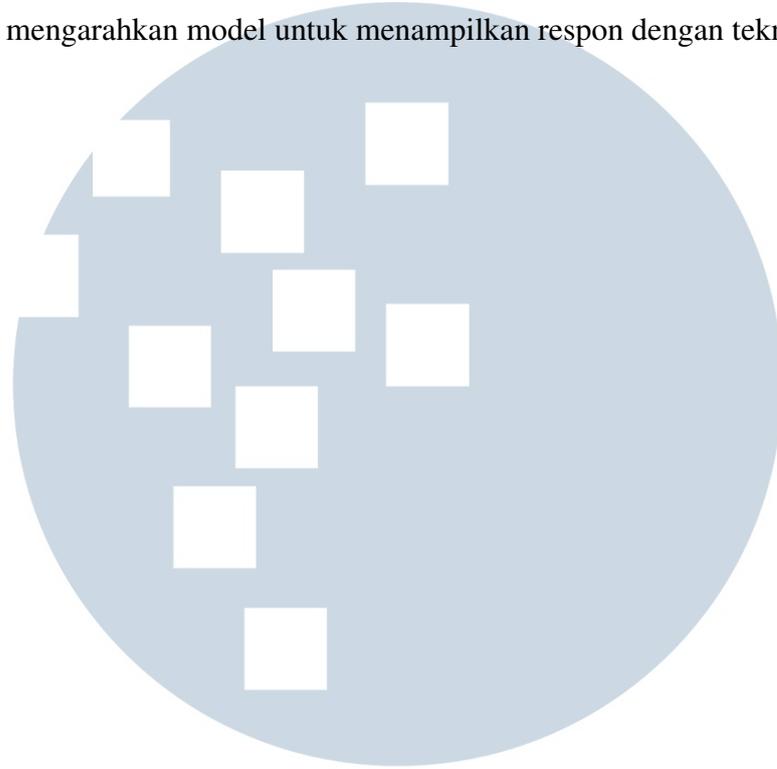
Terdapat beberapa kendala yang dialami selama pelaksanaan magang yang dilakukan di Badan Perencanaan Pembangunan Daerah Kabupaten Pohuwato, yaitu: **Kendala:**

1. Format data yang disimpan pada aplikasi *open data* Posdip cukup berantakan jika menyangkut data dengan subkolom pada tabelnya. Tidak API yang menyediakan data untuk mempermudah pengumpulan dan pemrosesan data. Kemudian pada tahap ini juga, server aplikasi Posdip masih dalam tahap *development* sehingga ketidak stabilan mengganggu proses pengumpulan dalam segi waktu yang cukup banyak.
2. Karena model harus dilatih dengan cloud, maka kemampuan jaringan internet diperlukan pada tahap ini. Kondisi dan tingkat infrastruktur daerah cukup menjadi kendala karena gangguan listrik maupun internet cukup umum terjadi dan menghambat proses pengerjaan.
3. Model Gemini tidak bisa bebas dikustomisasi seperti model chatbot lain yang *open source*. Penggunaan dan integrasi teknik lain pada model memerlukan akses berbayar. Maka integrasi dilakukan dengan sedikit kasar yang berdampak pula pada hasil akhir dari kemampuan model.

Solusi:

1. Menyampaikan kesalahan yang ada pada aplikasi Posdip untuk memperbaiki kualitas penyimpanan data yang ada. Kemudian memperbaiki data yang rusak sebelumnya dengan manual maupun membuat program pendukung untuk otomatis perbaikan dan peningkatan kualitas data.
2. Menggunakan konfigurasi yang lebih ringan pada pelatihan model. Mengurangi penggunaan sumber daya yang dibutuhkan dengan memperhatikan kemampuan model yang dikembangkan.
3. Menggunakan teknik lain, seperti RAG untuk mengurangi penggunaan sumber daya pada proses *fine tune*. Kemudian hal ini, juga membantu pada saat program dijalankan karena pencarian vektor tidak memerlukan internet dalam beroperasi.

4. Mempelajari pola respon jawaban ketika model tidak mengetahui data yang diminta. Membuat penanggulangan dengan pemeriksaan respon model yang akan mengarahkan model untuk menampilkan respon dengan teknik RAG.



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

BAB 4

SIMPULAN DAN SARAN

4.1 Simpulan

Studi *fine-tuning* model Gemini untuk proses analitik di Pemerintahan Pohuwato telah berhasil dibuat sesuai dengan *requirement* yang ditentukan. Perancangan Model Chatbot sebagai Asisten dalam Aplikasi Pos Satu Data Indonesia - Pohuwato telah berhasil dibuat menggunakan bahasa pemrograman Python dengan menggunakan library `google.generativeAi` untuk implementasi model. Model *chatbot* dilatih menggunakan data yang telah disediakan oleh berbagai OPD di Kabupaten Pohuwato lewat aplikasi Posdip. Spesifikasi data yang sudah integrasikan berjumlah 130 tabel mencakup bidang ekonomi, infrastruktur, kemiskinan, kependudukan, kesehatan, lingkungan hidup, pariwisata, pemerintahan, pendidikan, dan sosial. Model Chatbot ini sudah digunakan oleh pihak-pihak yang berkepentingan untuk membantu pekerjaan mereka di aplikasi *Open Data* Posdip.

4.2 Saran

Beberapa saran terhadap model Gemini untuk proses analitik di Pemerintahan Pohuwato adalah sebagai berikut:

1. Membuat fitur visualisasi plot otomatis lewat *prompt* dari pengguna untuk memberikan peran analisis yang lebih baik. Dengan begitu pekerjaan yang berhubungan dengan proses analisis akan lebih cepat dan mudah. Hal ini mempertimbangkan kebutuhan sumber daya manusia yang terbatas dalam pemerintahan Kabupaten Pohuwato.
2. Menambahkan kemampuan model untuk memproses bahasa daerah Gorontalo mengikuti karakteristik budaya dari pengguna. Mempermudah dan pengguna yang tidak terlalu fasih dalam menjelaskan bahasa yang baku dalam bahasa Indonesia. Memungkinkan masyarakat daerah mengakses model jika aplikasi *Posdip* ingin dibuka aksesnya untuk masyarakat luas.